

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:44:49 ; Search time 16.9714 Seconds  
(without alignments)  
2650.980 Million cell updates/sec

Title: US-09-497-967-7  
Perfect score: 2540  
Sequence: 1 MKNILVILIISLFNQIKS.....QCDPANELSILLISVYLL 468

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	775.5	30.5	395	2	A46031	immobilization sur
2	256.5	10.1	677	2	C42125	trophozoite cystei
3	250	9.8	1766	2	A42125	trophozoite cystei
4	236.5	9.3	667	2	A48579	trophozoite surfac
5	228.5	9.0	713	2	A35502	major surface-labe
6	228	9.0	1274	2	T42017	cysteine rich prot
7	218.5	8.6	1372	2	T25933	hypothetical prote
8	216.5	8.5	596	2	A45664	variant-specific s
9	197	7.8	557	2	A48434	proteoliasin - se
10	191.5	7.5	1297	2	T30274	hypothetical prote
11	190.5	7.5	2824	2	T22759	hypothetical prote
12	188.5	7.4	2823	2	T23064	protein T22A3.8 [i
13	188.5	7.4	2823	2	F87908	laminin alpha chain
14	188.5	7.4	3102	2	T43291	laminin gamma-1 ch
15	188	7.4	1609	1	MMHUB2	zonadhesin - mouse
16	188	7.4	5376	2	T42215	hypothetical prote
17	187.5	7.4	1459	2	T32271	furin (EC 3.4.21.7
18	183.5	7.2	1680	2	A43434	surface antigen se
19	181.5	7.1	439	2	A36385	serine proteinase
20	179.5	7.1	1299	2	T43251	laminin gamma-1 ch
21	179.5	7.1	1548	2	S34583	laminin alpha-1 ch
22	178.5	7.0	1607	1	MMMSB2	laminin alpha-1 ch
23	175.5	6.9	3075	2	S14458	G surface protein
24	174.5	6.9	3635	2	T10053	hypothetical prote
25	173	6.8	3712	2	S18253	hypothetical prote
26	172	6.8	2718	2	A23475	G surface protein
27	171	6.7	1895	2	T15881	hypothetical prote
28	169	6.7	1827	2	T34288	hypothetical prote
29	167.5	6.6	2704	2	S09118	G surface protein

30	167	6.6	2395	1	S50820	surface protein ty
31	166.5	6.6	969	1	A39490	subtilisin-like pr
32	166	6.5	3084	1	MMMSA	laminin alpha-1 ch
33	165	6.5	1700	2	S08167	Balbani ring 3 pr
34	164	6.5	1797	2	T21889	hypothetical prote
35	164	6.5	1805	2	T21888	hypothetical prote
36	160.5	6.3	7379	2	B88553	protein KOH4.2b [
37	160.5	6.3	2543	2	T31687	oocyst wall protei
38	159	6.3	1252	2	S36016	alpha-51D-immobil
39	158	6.2	2533	2	T28675	laminin alpha-2 ch
40	158	6.2	2533	2	T28674	hypothetical prote
41	158	6.2	3106	1	S53868	hypothetical prote
42	157	6.2	357	2	T21152	hypothetical prote
43	156	6.1	1391	2	T20406	hypothetical prote
44	156	6.1	1808	2	T15099	hypothetical prote
45	155.5	6.1	738	2	S40992	hypothetical prote

ALIGNMENTS

RESULT 1

A46031

immobilization surface I-antigen precursor - Ichthyophthirius multifiliis (fragment)  
C:Species: Ichthyophthirius multifiliis  
C>Date: 29-Sep-1999 #sequence\_revision 29-Sep-1999 #text\_change 07-Dec-1999  
C:Accession: A46031  
R:Clark, T.G.; McGraw, R.A.; Dickerson, H.W.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6363-6367, 1992  
A:Title: Developmental expression of surface antigen genes in the parasitic ciliate I  
A:Reference number: A46031; MUID:92335298; PMID:1631132  
A:Accession: A46031  
A:Molecule type: mRNA; protein  
A:Residues: 1-395 <CIA>  
A:Cross-references: GB:M92907; NID:g3628568; PIDN:AAC36158.1; PID:g3628569  
A:Note: the authors translated the codon UUG for residue 330 as Ile  
A:Note: sequence extracted from NCBI backbone (NCBIN:108734, NCBIP:108735); the sequ  
C:Genetics:  
A:Genetic code: SGC5  
C:Keywords: glycoprotein; surface antigen  
F:2-395/Product: immobilization surface I-antigen #status experimental <NAT>  
F:156,191,245,281/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.5%; Score 775.5; DB 2; Length 395;  
Best Local Similarity 44.2%; Pred. No. 3.7e+43;  
Matches 175; Conservative 35; Mismatches 129; Indels 57; Gaps 16;

QY	93	VKCPAGTAIAGGATDY-AAIITECVNCRINFY-----NENAP--NFNAGASTCTA	139
Db	2	VPCPDGTTQAGLTDVGAADLGTVCNCRPNFYNGGAOGEANGNPFAANNAARGICVP	61
QY	140	CPVNRVGGALTAGNAATIVACQNCVACPTGTALDDGTTDVTYRSTFCVRCRLNFYNGNN	199
Db	62	CQINRVGSVTNAGDLATLATQCTQCTGTALDDGTTDVTDFRSAAQCVKCPNFYNGGS	121
QY	200	--GNTP-----FNPG-----KSQCTPCPAKPAVNAQAATLGNDAITTAOCNVACPDG	244
Db	122	PQGEAPGVQVFAAGAAAGAAVTSQCVFCQINK--NDSPATAGAQAANLATQCSNOCPTG	179
QY	245	TISAAGVNNWVAQNT-CTNCAPNFYNN-----NAPN-----FNPG-----NSTC	283
Db	180	TVLDDGVT--LVFNTSATLCVKCRPNFYNGSGSQGEAPGVQVFAAGAAAGAAVTSQ	237
QY	284	LPCPANKDGAETAGTAGAATLAKQNCIACPDGTATASGAT-NYVILQTECLNCAANFYD	342
Db	238	VPQLAKN--DSPATAGAQAANLATQCSQCTGTATQDGTTLVFSNSSTQCSOCIANYFN	296
QY	343	GNNFOAGSSRCACAPANKVOGAVATAGGTATLQAACALCEPCAGTTLTGCTTSTYQAASE	402
Db	297	G-NLEAGSQCLKCPVSKTTPAHA-PGNTATQATCLTTCPCAGTDLDDGTSTNFVASATE	354
QY	403	CVKCAANFYTTKOTDHWAGIDTCTSCNKKLTSGAEA	438





Db 1003 -CVTCGYGYQP-----SAGAFECIPC-----GICKTTLSEFATSEDECRDECPDG 1047

QY 245 -TISAAGV-----NNWVAQNTCTNCAPNFYN-----NAPNFNP----- 279

Db 1048 EQLSAGVQPCOIGYRSGENKKVACPPGTTTAIMSTRREQNTPKCKPQGLVKE 1107

QY 280 NSTCLPCPKANDKYACATAGAGATLAKQCNIAOPDGTATIASGATNVVILQTECLN---CA 336

Db 1108 TKNQCPCPRGTFQNEQES-----TKLCAPDHTTAAPGAT-----AESCFSTNOCA 1155

QY 337 ANFYF-----DGNFQAGSSRCACPKANKVQAVATAGGTATLIAOALCECPAG 385

Db 1156 TGEYNCSWHANCIDLDPDENDVPSYECRCK--PGYRNGHTCTDA-----CNDFC--- 1202

QY 386 TVLTGDTTSTYKQAASECV-----KCAANFYTTKQTDWVAGI 422

Db 1203 --LNDGICKNNIGNVEICICKDHFSGDBRELRFQASNNKLIWATV 1245

RESULT 8

A45664

variant-specific surface protein VSP1267 - Giardia lamblia

C:Species: Giardia lamblia

C:Date: 22-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 29-Jan-1999

C:Accession: A45664

R:McMurtrei, M.R.; Aggarwal, A.; Nash, T.E.

Mol. Biochem. Parasitol. 49, 215-227, 1991

A:Title: Carboxy-terminal sequence conservation among variant-specific surface proteins of Giardia lamblia

A:Reference number: A45664; MUID:92131058; PMID:1775165

A:Accession: A45664

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-596 <MOW>

A:Cross-references: GB:M63966; NID:gl59140; PID:gl59141

A:Note: sequence extracted from NCBI backbone (NCBIN:77609, NCBI:77610)

Query Match 8.5%; Score 216.5; DB 2; Length 596;

Best Local Similarity 22.9%; Pred. No. 7.9e-07;

Matches 111; Conservative 40; Mismatches 180; Indels 153; Gaps 29;

QY 40 TPANVCNCKNFYNNAAAFVPGASTCTPCPKKADGAQNPATANLVTC----- 91

Db 117 TAGVCGACKDGYKNSDA--VATADSCIAC---EDANCATCGGAGENKCTCIDGYFVGA 171

QY 92 -----NVKCPAGTAIAGGATDYAAIITECVNCRINFYENAPNFNAGASTCTACPVNRV 145

Db 172 TGNEGCGIKCDATT-----GPNSYKG-VAGCAKC-----EKPK-NAGPAKCIETC----- 213

QY 146 GGALTAGNAATIVACQNVACPTGTALDDG-----VTTDYVRSFTE-CVKCRNLFFYNGNNG 200

Db 214 -----AADYLTKEADEQTSVSEAVCREGKTHFFPTDSAGGNKVCVSC----- 257

QY 201 NTFPNPKSQCTPCPAIKPANVAQATLGNDAITTA-----QCNVACPDGTISAA 249

Db 258 GTTNNGGIENGCECTSKESA---ARAGTEITCTKCSNNLSPLGDACLTDPCAGTAVS 313

QY 250 GVNWVA-----QNTCTNCAPNF-----YNNAPNFNPGNSTCL----- 284

Db 314 GDSGVCKPCNHTCAGCOTDRETCTACYPGYSLLYESNG-----ATGRCKECTGAFI 368

QY 285 -----PCPANKDYGAETAGGAATLAKQCNIAOPDGTATIA-----TAIASGATNVILQ 330

Db 369 TNCADGQCTAN-----VGS-----AKYCT-QCKDGYAPIDGICTAVAAAGRDVSVCTA 415

QY 331 -----ECLNCAANP-YPDGNNFOA-----GSSRCACPKANKVQAVAT-AGGTA-----TLIA 376

Db 416 TGGCTACTGNVALLSGGCYNTQTLPGKSVCKAV--ANSNDGCKCTCANGQAPDPATNECP 474

QY 377 QCALCEPAGTVLTDGTTST-----YKQAASECVKCAANFYTTKQTDWVAGIDTCTSCNKK 431

Db 475 LCDSTCAECSTKNDADACTCKEPPGYKTKNCKIKICTESSNNGKKID---GIPDCLSCAP 531

QY 432 LTSG 435

QY 21 ANCPV-GTETNAGQVLDLTPANVCNCKNFYNN-----NRAAFVPGASTCTPC 69

Db 801 APCNVGECCTVEGNAQ-----OCKTCRPGYTTINTDTKQCKDPEAPCNVGEGETCV-- 852

QY 70 POKKDAGAP---NPPATANLVQ-----CNVK-CPAGTAIAGGATDYAAIITEC 115

Db 853 -----EGNAQCKTCRPGYTTINTDTKQCKDPEAPCNVGECE--TCVEGNA-----QOC 899

QY 116 VNCNRYNFYENAPNFNAGASTCTACPVNRVGGALTAGNAATIVACQNVACPTGTALDDGV 175

Db 900 KTCRPGY-----TINTDTKQCKDPEAPCNV-----EAPCNV-----EGC 928

QY 176 TTDYVRSFTECVKRLNFYNGNNG-----NTPFNPKSQCTPCPAIKPANVAQA 225

Db 929 ETCVEGNAQCKTCRPGYTTINTDTKQCKDPEAPONT-----NCKTCDNPKTDNEI-C 981

QY 226 TLGNDATITACQNVACPDGTISAAV-----NNWVAQNTCTNCAPNFYNNAPNFNGN 280

Db 982 TKCNDGDYLTPTNOCVPDCT-AISGYIGDTKKCKACNPECAECV-----GPN 1029

QY 281 STCLPCPANK--DYGAETAGGAATLAKQCNIA-----CPD-----GTATASGATN--- 324

Db 1030 NQCTACPVGKMLQYTDNTNPVNGTGMDCQSVSTNDGCAEGCAIGGTAYCCKKNQ 1089

QY 325 -----YVILQTECLNCAANFYF-DG-----NNFOAGSSRC 353

Db 1090 APLNGNCAASSRVACATITSGACTKCEGYFLKDGCGYQTDROPKQVCSNAQGGNGKC 1149

QY 354 KACPANKVOGAVATAGGTATLIAQCALECPAGTVITDGTSTYKQAASECVKCAANFYTT 413

Db 1150 QTC-----ANGLAASDGNCA-ECHS-----TCATCST-ADAADKCKTCATGYIKE 1192

QY 414 KQTDWVAGIDTCTSCNKKLTSGAENLPESAKKNI 448

Db 1193 NGDDTTAGL--CKKCKSEKI-SGCKQCVSSGSSVI 1224

RESULT 7

T25933

hypothetical protein W02C12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T25933

R:Murray, J.; Wohldmann, P.

submitted to the EMBL Data Library, December 1996

A:Description: The sequence of C. elegans cosmid W02C12.

A:Reference number: Z20112

A:Accession: T25933

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1372 <MUR>

A:Cross-references: EMBL:U80815; PIDN:AAB37995.1; GSPDB:GN00022; CESP:W02C12.1

A:Experimental source: strain Bristol N2; clone W02C12

C:Genetics:

A:Gene: CESP:W02C12.1

A:Map position: 4

A:Introns: 29/1; 66/1; 774/2; 823/2; 1046/1; 1108/2; 1298/1

Query Match 8.6%; Score 218.5; DB 2; Length 1372;

Best Local Similarity 22.2%; Pred. No. 1.3e-06;

Matches 90; Conservative 35; Mismatches 155; Indels 125; Gaps 20;

QY 79 PN-PPATANLVQCNVKCPAGTAIAGGATDYAAIITECVNCRINFYENAPNFNAGASTC 137

Db 905 PNGRPDVGSLKIREYLCQAGVV-----VRDLVCPAPGTVH-----SAATGEC 949

QY 138 TACPVNR-----VGGALTAGNAATIVACQNVACPTGTALDDGVTIDYVRSFT 184

Db 950 ELCPIGEYQPLTARTECFKAPGQITASEGAISEGECKCNCPGHQY-DSLTS----- 1002

QY 185 ECVKCRNLFFYNGNNGTNPFPNPKSQCTPCPAIKPANVAQATLGNDAITTAQCNVACPDG 244

Db 532 INTG 535

## RESULT 9

A48434  
variant-specific surface protein - Giardia lamblia (strain GS/M)  
C:Species: Giardia lamblia  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-May-1999  
C:Accession: A48434  
R:Nash, T.E.; Mowatt, M.R.  
Mol. Biochem. Parasitol. 51, 219-228, 1992  
A:Title: Characterization of a Giardia lamblia variant-specific surface protein (VSP) gene  
A:Reference number: A48434; MUID:92244292; PMID:1574080  
A:Accession: A48434  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-557 <NAS>  
A:Cross-references: GB:M80480; NID:g159142; PID:g159143

Query Match 7.8%; Score 197; DB 2; Length 557;  
Best Local Similarity 20.7%; Pred. No. 1.4e-05;  
Matches 110; Conservative 58; Mismatches 211; Indels 152; Gaps 25;

QY 7 VILLISLFINOIKSANCPVGTETWAGVDDLGTPANCVCOKNFYNN----- 55  
Db 1 MFLINCLIASTLAGACSTTQANCAVEKCEMVGTEICTCKQNYVPINGVCEAAASNT 60  
QY 56 ----AAAFVPGASTCTPC-----PQKDKAGA----- 77  
Db 61 KQKASADEASDQTCGKCLSTTFWYKGCYDKTGNLGRITCKTEGSDAGKCGACNDEKGF 120  
QY 78 QPNPPATANLVTCNVK--PAGTATAGGATDYAAIITECVNCRINFINENAPNENAGAS 135  
Db 121 FDNPD--ANNVDS--ISCGDATGVTIPGSTKTKYGVAGCAK--TKPSOISENTGKEA 176  
QY 136 TCTACPNRVGGALTAGNATIAOQNV--CPTGTALDGVTTDYVRSFTECVKRLNF 193  
Db 177 TCTECNANLVKAVSSPTSAT---SCVSAEDCKTGYF---PTTDTTDDSKKCLTC--- 225  
QY 194 YNNGNNGTNPFGKSOCTPCPAIKPANVAQATL-----GNDATITAOQNVACPDGT 245  
Db 226 -----STADKGGIDCSACELLPTSTRASTVLISSACSTNNLSPLKNECMQDCPAGT 278  
QY 246 ISAAGV-----NWKVAQTE--CTNCAP-----NFYNNAPNF-----NPGN 280  
Db 279 YADSNVCKPCHTSCACKGDNTESSTACTYPGSVLSYGTDTKGTCTAECTGKYLENCAD 338  
QY 281 STCLPCPANKDYGAETAG----GAATLAKQCNACPDGTATAGTATAGTATVILQTECLNC 335  
Db 339 GOCTATIAGSKYCKSKSGFVNVGLCVSAETARAAPPST---PDKTNGV-----CTAC 390  
QY 336 AANFYFD-GNNFOA-----GSSRCKACAPK-----VOGAVATAGGTATLIAQCALECPAGT 386  
Db 391 TERYFLFSGCYQAEKFPFGNTLCTTADACKTTCANGQDKDSNGS-----CPA--- 438  
QY 387 VLTGDTSTYKQAASCVKCAANFY-----TTKQTDWAGIDTCSC 428  
Db 439 -CPTNASCACKONTKNTCNKCFSGYLDOTAKACKKCSSETSGNIQGVENCISC 488

## RESULT 10

T30274  
proteolisin - sea urchin (Strongylocentrotus purpuratus) (fragment)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C:Accession: T30274  
R:Laidlaw, M.; Wessel, G.M.  
Development 120, 1325-1333, 1994  
A:Title: Cortical granule biogenesis is active throughout oogenesis in sea urchins.  
A:Reference number: 220803; MUID:94298531; PMID:8026340  
A:Accession: T30274  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-1297 <LAI>  
A:Cross-references: EMBL:U57753; NID:g1373379; PID:g1373380; PIDN:AAB02256.1

Query Match 7.5%; Score 191.5; DB 2; Length 1297;  
Best Local Similarity 21.2%; Pred. No. 6.6e-05;  
Matches 125; Conservative 63; Mismatches 192; Indels 211; Gaps 37;

QY 21 ANCPVGTETNTAGQVDDLGTT-----PANCVNCOK 49  
Db 499 SGCPPTIVTCTPAGRI--DCGTNYCVVGARCDGVSDCSNGQDESGCPPTIVTCTPAGRIDCGT 557  
QY 50 NFYNNAAAFVPGASTCTPCPKKADAGAPNPATANLVTCNVKCPAGTATAGGATDYA 109  
Db 558 N--YCVVGARCDGVSDCS--NGQDESGCPP-----TIVT-----CPAGRIDCG--TNYC 600  
QY 110 AIIITEC---VNCRINFYNENAPNFNAGASTCTACPNRV---GGALTAGNATIAOQNV- 162  
Db 601 VVGARCDGVSDC-----SNGDEICGCPPTIVTCTPAGRVDCGNNYCVVSGKCDGVSDCSN 654  
QY 163 -----VACPTGTALDDGVTTDYVRSFTECV---KRLNFYINGNNGTTPENPGK 208  
Db 655 GODESGCPPTIVTCTPPG--RIDCG--TDY-----CVVGARC-----DGVSDCSNGO 696  
QY 209 SQ--CTP-----CPA-----IKPANVAQTALGNATITTAQCNVACPDGTISAA 249  
Db 697 DEIGCPPTIVTCTPAGRVDCGNNYCVVSGKCDGVSDCSNGQDESGCPPTTSACPEGRVDCG 756  
QY 250 GYNNVVAQNTC---TNCAPNFYNNAPNFNPGNSTCLPCPANK--DYGAETAGGAATLA 305  
Db 757 --NNYCVVSGKCDGVSDCSNGQDESGCP---PTTST---CPGVRVDCGTDYCVFGA--- 804  
QY 306 KCNTACPDGTATAGTATVILQTECLNCAANFYFDGNNFQAGSSRC----- 353  
Db 805 -RC-----DGVSDCSNGQDEIGCPPTIVTCTPAGRVDCGNNYCVVSGKCDGVSDCSNGODE 858  
QY 354 -----KACPANKVQ--GAVATAGTATLIAQCA-----LECPAGTV----- 387  
Db 859 SECPTTSCPEGRVDCGNNYCVVSGKCDGVSDCSNGQDESGCPPTIVTCTPAGRIDCGTN 918  
QY 388 -----LTDGTTT--TYKQAASEC-----YKCAANFYTT--KQTDWAGIDTC--- 425  
Db 919 YCVVGARCDGVSDCSNGQDESGCPPTIVTCTPAGRVDCGNNYCVVSGKCD---GVSDCSNG 975  
QY 426 -----TSCNKKLTS---GAEA--NLPESAKKNIQCFDANFLSISLLI 463  
Db 976 QDEEGCSFSSCRSGDCEFFGTEAVESLEQRVTKLEQLDLENILVLENILV 1026

## RESULT 11

T22759  
hypothetical protein F55H12.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Aug-2000  
C:Accession: T22759  
R:Dobson, R.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19610  
A:Accession: T22759  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2824 <WIL>  
A:Cross-references: EMBL:Z81091; PIDN:CAB03143.1; GSPDB:GN00019; CESP:F55H12.3  
A:Experimental source: clone F55H12  
C:Genetics:  
A:Gene: CESP:F55H12.3  
A:Map position: 1  
A:Introns: 87/1; 98/1; 126/2; 201/3; 343/3; 406/1; 576/3; 656/1; 825/3; 869/1; 909/1;  
7/1; 1755/2; 1800/1; 1850/3; 1896/1; 2003/3; 2035/3; 2082/3; 2119/1; 2144/1; 2200/2;  
C:Superfamily: LDL receptor ligand-binding repeat homology  
F;243-279/Domain: LDL receptor ligand-binding repeat homology <LDL>

Query Match 7.5%; Score 190.5; DB 2; Length 2824;



QY 99 TAIAGGATDYAALITECVNCRINFPNENAPNFNAGASTCTACPNVRVGGALTAGNAATV 158  
Db 964 ENYIG-----ACDRCK-----ENHGDVENG---CPACDNDTGS--IGSDCDQVS 1004  
QY 159 AQCNVACPTGTALDDGVTTDYVR---SFTF--CVKRLNFYNGNNGNTPFPNGKSOQ- 211  
Db 1005 GOCN--CKQGVF---GRQCDQCRPSYFNFTDAGCQFCNCHNIYGSIEDGKCDQTTGKCECR 1059  
QY 212 -----TPCPAIPKA--NVAQA-----TLGNDAT-----ITAQCNVACPDGTISAAGV 251  
Db 1060 ENVEGTWCKRCADGYENITSGDCGDCGCDPTGSDVSCNLVTGQC--VCKPG----- 1110  
QY 252 NNWVAQNTCTNCAPNFYNNAPNPNPNSGTCCLPCPANKDYGAETAGGAATLAKQCNTA 311  
Db 1111 -----VTGLKDCSCLPNFYGLT-----SEGCTECP-APGQVCDPDIAGSCV 1152  
QY 312 CPDGTATASGATNYVILQTECLNCAANF--YFDGNFNQAGSSRCKACPANKVQGVAVATAG 369  
Db 1153 CPPNT-----VGMCECNCCTTNAWDYHPLNG-----CKLDCDSD-----IGSDG 1190  
QY 370 GTA-TLIAQCALECPAGTAVLTGTTSTYKQAASECVKCAANFYTTKQTDWVAGIDTCTSC 428  
Db 1191 GMCNTFTGQC--KCKAAV-----GLKCDLCTHGFF-----NFTPTCEPC 1227

## RESULT 14

T43291

laminin alpha chain - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C:Accession: T43291

R:Zhu, X.; Kao, G.; Joh, K.; Hall, D.H.; Wadsworth, V.K.; Hutter, H.; Vogel, B.E.; Huang

submitted to the EMBL Data Library, June 1998

A:Description: Expression, function and evolution of laminin alpha chains.

A:Reference number: 22397

A:Accession: T43291

A&gt;Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-3102 &lt;2HU&gt;

A:Cross-references: EMBL:AF074902; PIDN:AAC26793.1

C:Genetics:

A:Map position: 1

A&gt;Note: lamal/2

C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match 7.4%; Score 188.5; DB 2; Length 3102;

Best Local Similarity 21.5%; Pred. No. 0.00023;

Matches 103; Conservative 32; Mismatches 160; Indels 185; Gaps 30;

QY 22 NCPVGTENTA-----GQVDDLGL-----GQVDDLGL-----TP 41

Db 860 NCPVGTENTA-----GQVDDLGL-----GQVDDLGL-----TP 41

QY 42 ANCVNCRNFYNNAAAFVPGASTCTPCPKQKADAGAPPPATANLVTC-----NVKCPAG 98

Db 920 DSCESC-KEHWNQAQ-----LHTCRPCGCHTQGVNPP-----QCSSENGEBCCK 963

QY 99 TAIAGGATDYAALITECVNCRINFPNENAPNFNAGASTCTACPNVRVGGALTAGNAATV 158

Db 964 ENYIG-----ACDRCK-----ENHGDVENG---CPACDNDTGS--IGSDCDQVS 1004

QY 159 AQCNVACPTGTALDDGVTTDYVR---SFTF--CVKRLNFYNGNNGNTPFPNGKSOQ- 211

Db 1005 GOCN--CKQGVF---GRQCDQCRPSYFNFTDAGCQFCNCHNIYGSIEDGKCDQTTGKCECR 1059

QY 212 -----TPCPAIPKA--NVAQA-----TLGNDAT-----ITAQCNVACPDGTISAAGV 251

Db 1060 ENVEGTWCKRCADGYENITSGDCGDCGCDPTGSDVSCNLVTGQC--VCKPG----- 1110

QY 252 NNWVAQNTCTNCAPNFYNNAPNPNPNSGTCCLPCPANKDYGAETAGGAATLAKQCNTA 311

Db 1111 -----VTGLKDCSCLPNFYGLT-----SEGCTECP-APGQVCDPDIAGSCV 1152

QY 312 CPDGTATASGATNYVILQTECLNCAANF--YFDGNFNQAGSSRCKACPANKVQGVAVATAG 369  
Db 1153 CPPNT-----VGMCECNCCTTNAWDYHPLNG-----CKLDCDSD-----IGSDG 1190  
QY 370 GTA-TLIAQCALECPAGTAVLTGTTSTYKQAASECVKCAANFYTTKQTDWVAGIDTCTSC 428  
Db 1191 GMCNTFTGQC--KCKAAV-----GLKCDLCTHGFF-----NFTPTCEPC 1227

## RESULT 15

MMHUB2

laminin gamma-1 chain precursor - human

N:Alternate names: laminin chain B2

C:Species: Homo sapiens (man)

C&gt;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 10-Dec-1999

C:Accession: S13548; A28158; S13549; S14664; S23567

R:Kallunki, T.; Ikonen, J.; Chow, L.T.; Kallunki, P.; Tryggvason, K.

J. Biol. Chem. 266, 221-228, 1991

A:Title: Structure of the human laminin B2 chain gene reveals extensive divergence f

A:Reference number: S13548; MUID:91093128; PMID:1985895

A:Accession: S13548

A&gt;Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1609 &lt;KAL&gt;

A:Cross-references: GB:M55217; NID:g186937

A&gt;Note: the nucleotide sequence was submitted to GenBank, February 1991

R:Pikkarainen, T.; Kallunki, T.; Tryggvason, K.

J. Biol. Chem. 263, 6751-6758, 1988

A:Title: Human laminin B2 chain. Comparison of the complete amino acid sequence with

A:Reference number: A28158; MUID:88198245; PMID:3360804

A:Accession: A28158

A:Molecule type: mRNA

A:Residues: 1-2111, 'I', 213-1609 &lt;PIK&gt;

A:Cross-references: EMBL:J03202; NID:g186916; PIDN:AAA59488.1; PID:g307107

R:Fukushima, Y.; Pikkarainen, T.; Kallunki, T.; Eddy, R.L.; Byers, M.G.; Haley, L.L.;

Cytogenet. Cell Genet. 48, 137-141, 1988

A:Title: Isolation of a human laminin B2 (LAMB2) cDNA clone and assignment of the gen

A:Reference number: S13549; MUID:89169663; PMID:3234037

A:Accession: S13549

A:Molecule type: mRNA

A:Residues: 1393-1609 &lt;FUK&gt;

A:Cross-references: EMBL:M27654; NID:g186923; PIDN:AAA59489.1; PID:g186924

R:Olsson, D.; Nagayoshi, T.; Fazio, M.; Peltonen, J.; Jaakkola, S.; Sanborn, D.; Sasak

Lab. Invest. 60, 772-782, 1989

A:Title: Human laminin: cloning and sequence analysis of cDNAs encoding A, B1 and B2

A:Reference number: A34961; MUID:89280632; PMID:2733383

A:Accession: B34961

A:Molecule type: mRNA

A:Residues: 868-1551, 'N', 1553-1609 &lt;OLS&gt;

R:Santos, C.L.S.; Sabbaga, J.; Brentani, R.

DNA Seq. 1, 275-277, 1991

A:Title: Differences in human laminin B2 sequences.

A:Reference number: S14664; MUID:92216129; PMID:1806043

A:Accession: S14664

A:Molecule type: mRNA

A:Residues: 1282-1609 &lt;SAN&gt;

A:Cross-references: EMBL:X13939; NID:g34237; PIDN:CAA32122.1; PID:g34238

R:Vuolteenaho, R.; Kallunki, T.; Chow, L.; Ikonen, J.; Pikkarainen, T.; Tryggvason, K

in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp. 175-193, Academi

A:Title: Genes for the human laminin B1 and B2 chains.

A:Reference number: S23566

A:Accession: S23567

A:Molecule type: DNA

A:Residues: 801-1481, 'R', 1483-1609 &lt;VUU&gt;

A&gt;Note: mRNA was also sequenced

C:Genetics:

A:Gene: GDB:LAMC1; LAMB2

A:Cross-references: GDB:120136; OMIM:150290

A:Map position: lq31-lq31

A:Introns: 140/1; 241/3; 285/2; 341/1; 404/1; 443/2; 476/2; 522/1; 563/1; 626/2; 664/

/3; 1525/1

C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type lamin

C:Function:

A>Description: interact with cells and with other basement membrane proteins to promote C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular F:1-33/Domain: signal sequence #status predicted <SIG> F:34-1609/Product: laminin gamma-1 chain #status predicted <MAT> F:34-285/Domain: VI <DOM6> F:34-285/Domain: VI <DOM6> F:286-504/Domain: V <DOM5> F:286-339/Domain: laminin-type EGF-like homology <LE01> F:286-339/Domain: laminin-type EGF-like homology <LE02> F:342-395/Domain: laminin-type EGF-like homology <LE03> F:398-442/Domain: laminin-type EGF-like homology <LE04> F:445-492/Domain: laminin-type EGF-like homology <LE05> F:495-504/Domain: laminin-type EGF-like homology #status atypical <LE05> F:505-689/Domain: laminin-type EGF-like homology <LE06> F:690-1034/Domain: III <DOM3> F:690-1034/Domain: laminin-type EGF-like homology #status atypical <LE06> F:724-770/Domain: laminin-type EGF-like homology <LE07> F:724-770/Domain: laminin-type EGF-like homology <LE08> F:773-825/Domain: laminin-type EGF-like homology <LE09> F:828-881/Domain: laminin-type EGF-like homology <LE10> F:884-932/Domain: laminin-type EGF-like homology <LE11> F:935-980/Domain: laminin-type EGF-like homology <LE12> F:983-1028/Domain: laminin-type EGF-like homology <LE12> F:1035-1609/Domain: II/I <DOM1> F:1035-1609/Region: heptad repeats F:40-50/Disulfide bonds: #status predicted F:60-134,576,650-1022,1107,1161,1175,1205,1223,1241,1380,1395,1439/Binding site: carbonyl F:1031,1034,1600/Disulfide bonds: interchain #status predicted

Query Match	7.4%	Score 188	DB 1	Length 1609
Best Local Similarity	23.9%	Pred. No. 0.00014		
Matches 99	Conservative	Mismatches 165	Indels 122	Gaps 26
QY	75	AGAAPNPATANLVTCQNVKCPAGTAIAGGATDYAAITTEVCNCRINFINYNAPNFNAGA	134	
DB	674	ASARFGPGVPATWVESCT--CPVG--YGGQF-----CEMC-LSGYRRETPNLGP-Y	718	
QY	135	STCTACPVNRVGGALTAGNAATIAQ--NVACPTGTALDDGVTTDYVRSFTECVKCRLN	192	
DB	719	SPCVILCACN---GHSETCDPETGVNCRDNTAGP-----HCEKCSDG	757	
QY	193	FYVNGNNGTTPFPKSKQCTPCP-----ATKP-----ANVAQTILGNDAITACN	238	
DB	758	YYGDSLTAGTS-----SDQPCPCPGSSCAVVPKTEVVCNCPGTGK-----RCE	805	
QY	239	VACPDGTISAAGVNNVVAQNTCTNCAPNFNNAPNFPNGNSTCLPCPAN-----	289	
DB	806	L-CDDGTFGDPILGRNGPVLRLCQCSDNIDPNAVGNCNRLTGECLKCIYNTAGFYCDRC	864	
QY	290	KD-YGAETAGGAATLAKQCNIA-C-PDGTASAGATNYVLQTECL-----NCAAN	338	
DB	865	KDGFNGPLAPNPADKCKACN--CNFYGTMKQSSCNPNVQGCECLPHVTGQDCGACDPC	922	
QY	339	FYEDGNNGFOAGSSRCKACPANKVOGAVATAGGTATL-IAQCALECPAGTVLTDGTTSTYK	397	
DB	923	FY-----NLQSGQG-CERCDCH---ALGSTNGQCDINTGQC--ECOPI-----	960	
QY	398	QAASECVKCAANFYTTKQTDWAGIDTCTSCNKKLTISGAEANLPESAKKNIQCD	451	
DB	961	-TGHCERCEVNHFG-----FGPECKPCD-----CHPEGSLQCKDGRCE	1002	

Search completed: February 11, 2003, 19:48:47  
Job time : 24.9714 secs